

TUFTS UNIVERSITY
Medical School

Prof. M. Bianconi
Course CMBA0264-01; Statistics with Applications; Summer 2009
E-Mail: marcelo.bianconi@tufts.edu

Assignment I

Data sets and other materials posted at the course web page:
<http://www.tufts.edu/~mbiancon/CMBA0264-2009.html>

Due date: Tuesday, June 30, 2009 (by 6:00PM via e-mail)

1. Please, write an executive summary of your reading of chapters 3 and 4 of Ian Ayres' *Super Crunchers*.
2. Please, write an executive summary of your reading of chapters 15 of Nassim Taleb's *Black Swan* (Available at the course web page).
3. A coin is tossed three times. What is the probability that it lands on heads *only* one time?
4. "The hitting coach, former Red Sox slugger Jim Rice, rode [Scott Hatteberg] long and hard. Rice called him out in the clubhouse, in front of his teammates, and ridiculed him for having a batting average in the .270s when he hit .500 when he swung at the first pitch. "Jim Rice hit like a genetic freak and he wanted everyone else to hit the way he did," Hatteberg said. "He didn't understand that the reason I hit .500 when I swung at the first pitch was that I only swung at first pitches that were too good not to swing at."
What's your statistics interpretation of the comments above?

INTRODUCTION TO STATA

0. Open the STATA program and go to:
File > Example Datasets > Example Datasets Installed with STATA > citytemp.dta use

Variables names and labels should appear in Variables.
Type: Describe

1. Report the output (copy text and paste) and describe the data and their type: categorical, numerical, discrete, continuous...
2. Construct a table of the frequency distribution of the variable **region** with the command: tabulate region

Copy and paste, and write a one line comment.

3. Construct a bar chart for the count of **heatdd** over regions using the command: `graph bar (count) heatdd, over(region)`

In the Edit menu, copy and paste the graph into your answer. Write a one line comment.

4. Construct a pie chart for the **heatdd** over regions: `graph pie heatdd, over(region)`

In the Edit menu, copy and paste the graph into your answer. Write a one line comment.

5. Construct a graph of the frequency histogram of the variable **heatdd**: `histogram heatdd, freq`

Copy and paste, and write a one line comment.

6. Construct a bar chart for the mean of **tempjan** over regions:

`graph bar (mean) tempjan, over(region)`

In the Edit menu, copy and paste the graph into your answer. Write a one line comment.

7. Construct a bar chart for the mean of **tempjuly** over regions:

`graph bar (mean) tempjuly, over(region)`

In the Edit menu, copy and paste the graph into your answer. Write a one line comment and compare with the answer in 6.

8. Construct graphs of the frequency histogram of the variables **tempjan** and **tempjuly**:

`histogram tempjan, freq`

`histogram tempjuly, freq`

Copy and paste, and write a one line comment each and compare.